Localized Active Learning of Gaussian Process State Space Models


Alexandre Capone
alexandre.capone@tum.de


Gerrit Noske
gerrit@noske.info


Jonas Umlauft
jonas.umlauft@tum.de


Thomas Beckers
t.beckers@tum.de


Armin Lederer
armin.lederer@tum.de


Sandra Hirche
hirche@tum.de

Motivation


[robohub.org]

- Accuracy of learned dynamic models depends on collected data
  → how to explore system efficiently?
- Exploration strategies often aim to obtain a globally accurate model [1]
- Control laws often only require a **locally accurate model** to perform well

Idea

Exploring only region of interest increases exploration efficiency

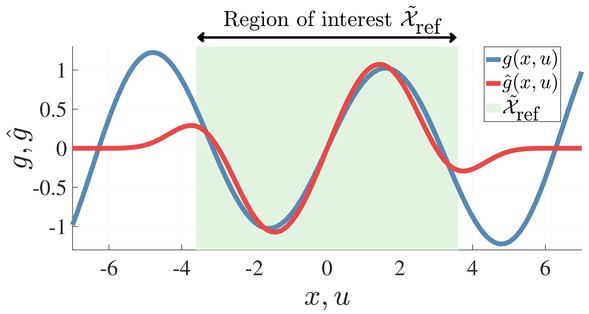
Problem Setting

Consider a partially unknown system $\boldsymbol{x}_{t+1} = \underbrace{\boldsymbol{f}(\boldsymbol{x}_t, \boldsymbol{u}_t)}_{\text{known}} + \underbrace{\boldsymbol{g}(\boldsymbol{x}_t, \boldsymbol{u}_t)}_{\text{unknown}} + \underbrace{\boldsymbol{w}_t}_{\text{process noise}}$.

Goal

Explore dynamics to obtain a **locally accurate** model $\widehat{\boldsymbol{g}}(\cdot)$ of $\boldsymbol{g}(\cdot)$.

$\rightarrow \widehat{\boldsymbol{g}}(\boldsymbol{x}, \boldsymbol{u}) \approx \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{u})$ for all $[\boldsymbol{x}^\mathsf{T}, \boldsymbol{u}^\mathsf{T}]^\mathsf{T} \in \tilde{\mathcal{X}}_{\text{ref}}$, where $\tilde{\mathcal{X}}_{\text{ref}}$ is a bounded set.



Gaussian Processes

The unknown function $\boldsymbol{g}(\cdot,\cdot)$ is modeled as a normal distribution

$$\boldsymbol{g}(\tilde{\boldsymbol{x}}_{t+1}) \sim \mathcal{N}\left(\boldsymbol{\mu}(\tilde{\boldsymbol{x}}_t), \boldsymbol{\Sigma}^2(\tilde{\boldsymbol{x}}_t)\right),$$

where $\boldsymbol{\mu}(\tilde{\boldsymbol{x}}_t)$, $\boldsymbol{\Sigma}^2(\tilde{\boldsymbol{x}}_t)$ depend on collected data and covariance function $k(\cdot,\cdot)$.

Interpretation

$\boldsymbol{\mu}(\tilde{\boldsymbol{x}}_t)$ is the **estimated model**, $\boldsymbol{\Sigma}^2(\tilde{\boldsymbol{x}}_t)$ represents **model uncertainty** [2].

Quantifying Utility of Data

Given a **finite discretization** $\tilde{\boldsymbol{X}}_{\text{ref}}$ of $\tilde{\mathcal{X}}_{\text{B}}$, the **mutual information** of a data point $\boldsymbol{\xi} \in \tilde{\mathcal{X}}$ with respect to $\tilde{\boldsymbol{X}}_{\text{ref}}$ is given by

$$I(\boldsymbol{y}_{\boldsymbol{\xi}}, \boldsymbol{y}_{\tilde{\boldsymbol{X}}_{\text{ref}}}) = \frac{1}{2}\log\left(\frac{(k(\boldsymbol{\xi},\boldsymbol{\xi})+\sigma)|\boldsymbol{K}_{\tilde{\boldsymbol{X}}_{\text{ref}}}+\sigma\boldsymbol{I}|}{|\boldsymbol{K}_{\tilde{\boldsymbol{X}}_{\text{ref}}\cup\boldsymbol{\xi}}+\sigma\boldsymbol{I}|}\right).$$

Exploration Algorithm

Algorithm 1 LocAL (Localized Active Learning)

Input: $\boldsymbol{x}_0$, $\boldsymbol{f}(\cdot)$, $k(\cdot,\cdot)$

1: **for** $t = 0, 1, 2, 3, \ldots$ **do**

2: Compute **point of maximal information** $\boldsymbol{\xi}^* = \arg\max_{\boldsymbol{\xi}\in\tilde{\mathcal{X}}} I(\boldsymbol{y}_{\boldsymbol{\xi}}, \boldsymbol{y}_{\tilde{\boldsymbol{X}}_{\text{ref}}})$

3: Compute **solution of optimal control problem**

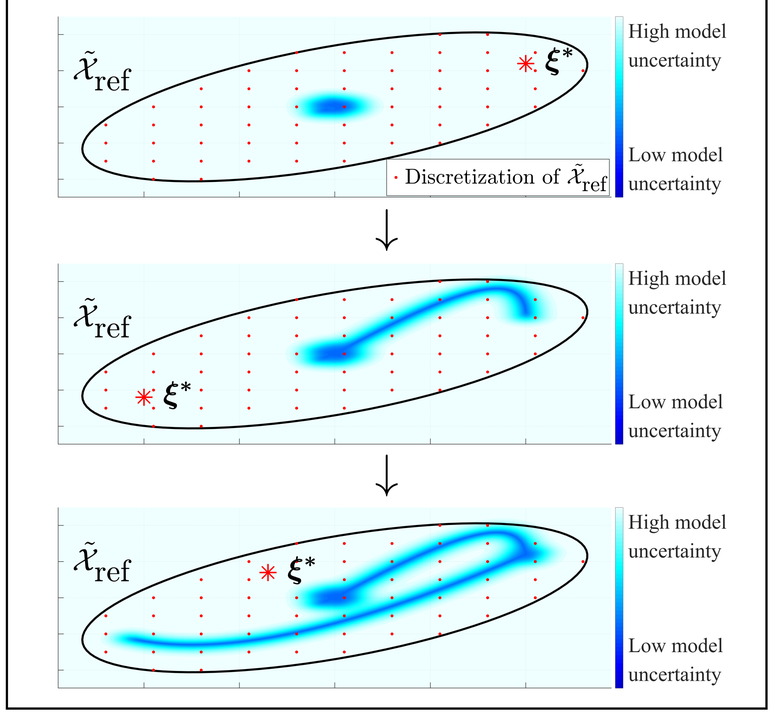
$$\boldsymbol{U}^* = \arg\min_{\boldsymbol{U}\in\mathcal{U}^{N_H}} \sum_{t=1}^{N_H} (\boldsymbol{\xi}^* - \tilde{\boldsymbol{x}}_t)^\mathsf{T} \boldsymbol{Q} (\boldsymbol{\xi}^* - \tilde{\boldsymbol{x}}_t)$$

$$\text{s.t.} \quad \boldsymbol{x}_{t+\tau+1} = \boldsymbol{f}(\tilde{\boldsymbol{x}}_{t+\tau}) + \boldsymbol{\mu}_t(\tilde{\boldsymbol{x}}_{t+\tau}), \quad \boldsymbol{u}_{t+\tau} \in \mathcal{U}, \quad \forall \tau \in \{0, \ldots, N_H - 1\}$$

4: Apply $\boldsymbol{u}^*_{t+1}$ to system, measure $\boldsymbol{x}_{t+1}$, update GP model $\boldsymbol{\mu}_t(\cdot)$, $\boldsymbol{\Sigma}^2_t(\cdot)$

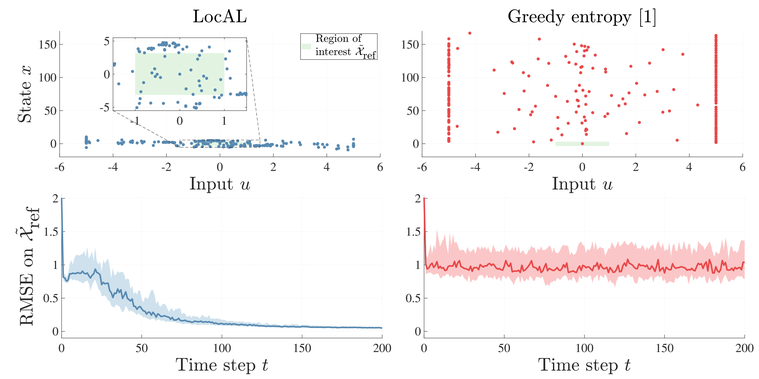
5: **end for**

Visualization

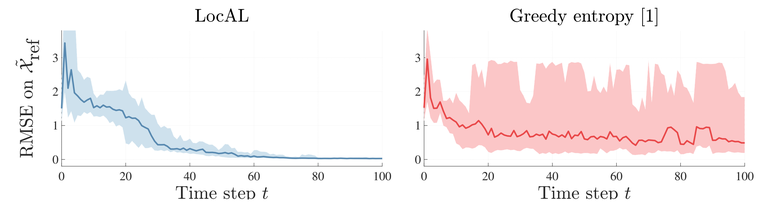


Monte Carlo Simulations

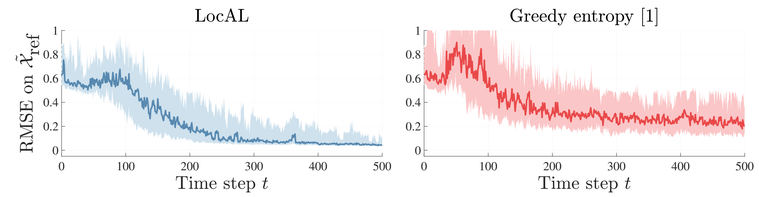
Toy problem



Pendulum



Cart-pole



Comparison

Our method **consistently yields a more accurate model** on $\tilde{\mathcal{X}}_{\text{ref}}$ than the greedy entropy-based approach from [1].

References

[1] J. Schreiter, D. Nguyen-Tuong, M. Eberts, B. Bischoff, H. Markert, and M. Toussaint, "Safe exploration for active learning with Gaussian processes," in *Machine Learning and Knowledge Discovery in Databases*. Springer, 2015, pp. 133–149.

[2] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, Jan. 2006.

Technische Universität München
Chair of Information-oriented Control